J Arid Land (2023) 15(6): 710–723
https://doi.org/10.1007/s40333-023-0103-7

Science Press Springer-Verlag

Modelling the dead fuel moisture content in a grassland of Ergun City, China

CHANG Chang[1,2], CHANG Yu[1*], GUO Meng[3], HU Yuanman[1,4]

[1] CAS Key Laboratory of Forest Ecology and Management, Institute of Applied Ecology, Chinese Academy of Sciences, Shenyang 110016, China;
[2] University of Chinese Academy of Sciences, Beijing 100049, China;
[3] School of Geographical Sciences, Northeast Normal University, Changchun 130024, China;
[4] E'erguna Wetland Ecosystem National Research Station, Hulunbuir 022250, China

Abstract: The dead fuel moisture content (DFMC) is the key driver leading to fire occurrence. Accurately estimating the DFMC could help identify locations facing fire risks, prioritise areas for fire monitoring, and facilitate timely deployment of fire-suppression resources. In this study, the DFMC and environmental variables, including air temperature, relative humidity, wind speed, solar radiation, rainfall, atmospheric pressure, soil temperature, and soil humidity, were simultaneously measured in a grassland of Ergun City, Inner Mongolia Autonomous Region of China in 2021. We chose three regression models, i.e., random forest (RF) model, extreme gradient boosting (XGB) model, and boosted regression tree (BRT) model, to model the seasonal DFMC according to the data collected. To ensure accuracy, we added time-lag variables of 3 d to the models. The results showed that the RF model had the best fitting effect with an R^2 value of 0.847 and a prediction accuracy with a mean absolute error score of 4.764% among the three models. The accuracies of the models in spring and autumn were higher than those in the other two seasons. In addition, different seasons had different key influencing factors, and the degree of influence of these factors on the DFMC changed with time lags. Moreover, time-lag variables within 44 h clearly improved the fitting effect and prediction accuracy, indicating that environmental conditions within approximately 48 h greatly influence the DFMC. This study highlights the importance of considering 48 h time-lagged variables when predicting the DFMC of grassland fuels and mapping grassland fire risks based on the DFMC to help locate high-priority areas for grassland fire monitoring and prevention.

Keywords: dead fuel moisture content (DFMC); random forest (RF) model; extreme gradient boosting (XGB) model; boosted regression tree (BRT) model; grassland; Ergun City

Citation: CHANG Chang, CHANG Yu, GUO Meng, HU Yuanman. 2023. Modelling the dead fuel moisture content in a grassland of Ergun City, China. Journal of Arid Land, 15(6): 710–723. https://doi.org/10.1007/s40333-023-0103-7

1 Introduction

Grasslands, which cover one-third of the Earth's terrestrial surface, are the second largest terrestrial carbon sink and host high levels of biodiversity (Lee et al., 2020; Petermann and Buzhdygan, 2021; Muro et al., 2022). Grassland fires are an integral part of grassland ecosystems worldwide, and in some circumstances, they can increase grassland biodiversity (Deak et al., 2014) and improve the performance of grazing livestock (Limb et al., 2011). However, grassland fires also produce large amounts of greenhouse gases and cause major economic losses to human society (Yebra et al., 2008; Sharma et al., 2021). Grassland fires are also sudden and highly

*Corresponding author: CHANG Yu (E-mail: changyu@iae.ac.cn)
Received 2022-10-21; revised 2023-02-11; accepted 2023-02-27

destructive disasters that can not only alter the structure, function, pattern, and processes of the landscape but also pose threats to herders' lives, infrastructure, and valuable grassland resources (Podur et al., 2003; Fontenele et al., 2020). The dead fuel moisture content (DFMC) has been regarded as a significant determinant of wildfire risk (Fernandes, 2001; Dragozi et al., 2021) because it affects the ignition risk (Wilson, 1985) and combustion rate (Catchpole et al., 1998; Hiers et al., 2019).

However, in previous studies on the DFMC, researchers have focused on forest fuels (Bakšić et al., 2017; Lee et al., 2020; Resco de Dios et al., 2021). Research on the DFMC of grassland fuels is generally neglected. In fact, grassland and savanna wildfires are particularly widespread, accounting for approximately 90% of the global area burned in the last century (Mouillot and Field, 2005). These grassland wildfires were large and devastating (Sharma et al., 2021). According to statistics, from 1980 to 2018, there were more than 1200 grassland fires recorded in the Inner Mongolia Autonomous Region of China, resulting in a loss of 5.99×10^7 CNY and 29 deaths, greatly influencing the lives of local herders. Therefore, accurately estimating the DFMC will be beneficial for predicting the occurrence probability of grassland fires to assist relevant departments in the prevention and combat of grassland fires in a timely and effective manner (Schunk et al., 2017).

Numerous researchers have built the DFMC prediction models that can be divided into two categories: process-based models and empirical models (Matthews et al., 2010; Sun et al., 2021). Process-based models, for example, the Simard model, Van Wagner model, Anderson model, and Nelson model (Nelson, 2000), are used to simulate vapour exchange in the interior of the DFMC based on the time-lagged equilibrium moisture content (Matthews, 2014). These kinds of models can predict the DFMC exactly, so the forest fire weather index of Canada and the National Fire Danger Rating System of the USA both adopt the equilibrium moisture content model (Stocks et al., 1989). However, due to the different physicochemical properties of various fuels, the response process of moisture variation will be different; thus, the parameters of process-based models need to be adjusted in other studies (Sun et al., 2021). Empirical models use statistical linear regression to determine the relationship between measured moisture content data and meteorological factors (Resco de Dios et al., 2015). As they do not consider the physicochemical properties of fuels, empirical models are easier to use than process-based models. Among the numerous empirical models, the multiple linear regression (MLR) model based on meteorological factors is the most basic and common prediction model (Bilgili et al., 2018; Man et al., 2019) and is often regarded as a benchmark (Shmuel et al., 2022). However, empirical models always need more observational data to ensure their accuracy, just as the prediction accuracies of some MLR models cannot meet the needs of the studies. Thus, in recent years, machine learning models have been increasingly favoured by researchers for their high predictive power and fast calculation speed. Although machine learning models are a kind of empirical model, they have been proven to have higher accuracy in fuel moisture content prediction in many studies (Lee et al., 2020; Zhu et al., 2021; Cunill Camprubí et al., 2022). For example, Capps et al. (2021) used the random forest (RF) model to estimate the live fuel moisture content (LFMC) in California, USA. Lei et al. (2022) used homemade monitoring equipment to obtain the DFMC value and weather conditions and built a prediction model using a backpropagation neural network, which is a type of machine learning method. In addition, mixed effects models (Xing and Qu, 2017), generalised additive models (Masinda et al., 2021), and time series prediction models (Fan and He, 2021) were also used to predict the DFMC with good prediction results.

However, most of these empirical models depend on real-time meteorological variables only and ignore the effect of time lags. The absolute change in the moisture content of fuels after different time lags may be different, so time lags are key factors influencing the DFMC of grassland. Only a few studies have added time-lag variables into empirical models, such as the time-lag variables of rainfall (Jin and Li, 2014) and relative air humidity (Zhang et al., 2015), which were occasionally added into modelling. In fact, fuels need a certain period of time to reach

the equilibrium moisture content due to their different sizes. Past equilibrium moisture content values and meteorological conditions could affect the current DFMC values (Shmuel et al., 2022). In addition, soil humidity (Rakhmatulina et al., 2021), soil temperature (Sun et al., 2021), air temperature (Nieto et al., 2010), wind speed, radiation (Masinda et al., 2021), and air pressure also have effects on the DFMC. Therefore, adding several environmental variables and their derived time-lag variables into the prediction model is likely to improve the model prediction accuracy for the DFMC.

The grassland of Ergun City is located in the forest-grassland transition zone in northern China, which is sensitive to climate. Under the condition of global climate change, the variation rules of the DFMC in the grassland of Ergun City will be more complex. Therefore, in this study, we used three regression models to build the DFMC prediction models for the grassland of Ergun City based on the measured DFMC values and related environmental variables, as well as their time-lag variables. Our aims were to determine the most suitable prediction model for the DFMC of grassland, explore whether the DFMC value has different influencing factors in each season, and determine the most appropriate length of the added time-lag variables, so as to more accurately and effectively predict the occurrence of grassland fires and optimize fire prevention measures in different seasons.

2 Materials and methods

2.1 Study area

The study area is in the northeastern Inner Mongolia Autonomous Region of China, which has a temperate continental monsoon climate. The grassland type here is an upland meadow, and the eastern part of the sampling site is near cultivated land (Fig. 1). The annual average temperature is –2.4°C, and the average annual precipitation is approximately 361.6 mm, most of which is concentrated in summer, with the least amount of precipitation falling in winter (Di et al., 2019). It is located in a transition zone between forest and grassland in northern farming-pastoral transitional zone of China (Di et al., 2019).

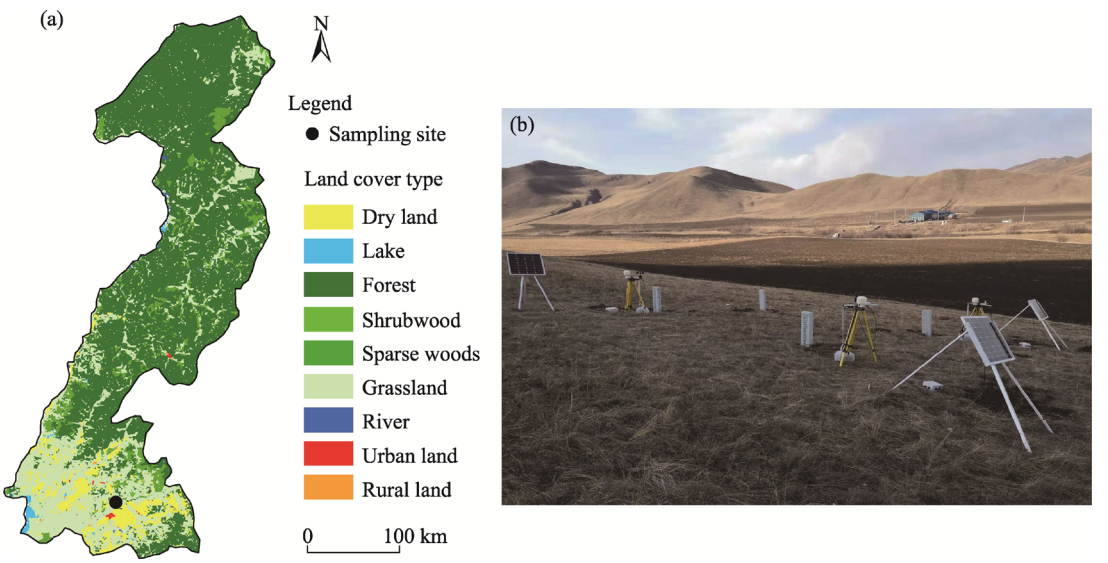

Fig. 1 Land cover types of Ergun City (a) and an overview of fuel moisture content meters (b). The land cover type data are derived from https://www.resdc.cn/Default.aspx.

The sampling site (50°19′45″N, 120°13′49″E) is located in the south of Ergun City, situated on a sunny slope at an altitude of 622 m (Fig. 1). At the sampling site, *Leymus chinensis* (Trin.) Tzvel. is the dominant species; in addition, *Saposhnikovia divaricata* (Trucz.) Schischk. and

Pulsatilla turczaninovii Kryl. et Serg. are distributed sporadically. In recent years, with the development of the tourism, an increasing number of tourists have travelled to Ergun City. The increase in the number of tourists also introduces certain fire risks to the local environment.

2.2 Dead fuel moisture content (DFMC) monitoring method

To obtain real-time and long-term data on meteorological conditions and fuel moisture, we deployed three fuel moisture content meters (FMC-M3, Northeast Forestry University, Harbin, China) at the sampling site (Fig. 2). The meters can automatically weigh samples and obtain meteorological data at the moment of weighing. Users can set the sampling intervals as needed. Solar panels can continuously power the batteries of the meters to keep them working without supervision. In order to check the accuracy of each meter balance, we randomly weighed a certain mass of fuel using an ordinary electronic scale and then compared it with the weight obtained from the meter balance. The error is less than 0.01 g (Masinda et al., 2021). We placed samples of dead herbaceous plants from the sampling site into 3 mesh bags, transported them to the laboratory, and dried the samples in an oven at 100°C for 24 h. After drying, the samples were weighed to obtain the dry weight. Next, we returned them to the sampling site and tied them under the weighted levers (an automatic balance) of the three meters. When the weighted levers, which were set to a specific interval, were triggered, the mesh bags were lifted; at all other times, the mesh bags remained on the ground. The meters can automatically measure fuel weight, air temperature, relative air humidity, wind speed, and solar radiation (Masinda et al., 2022), and all data can be transmitted via Bluetooth to a smartphone using the appropriate application. To ensure measurement accuracy, the meters stop working when the wind speed is over 3 m/s or when the temperature is below 5°C.

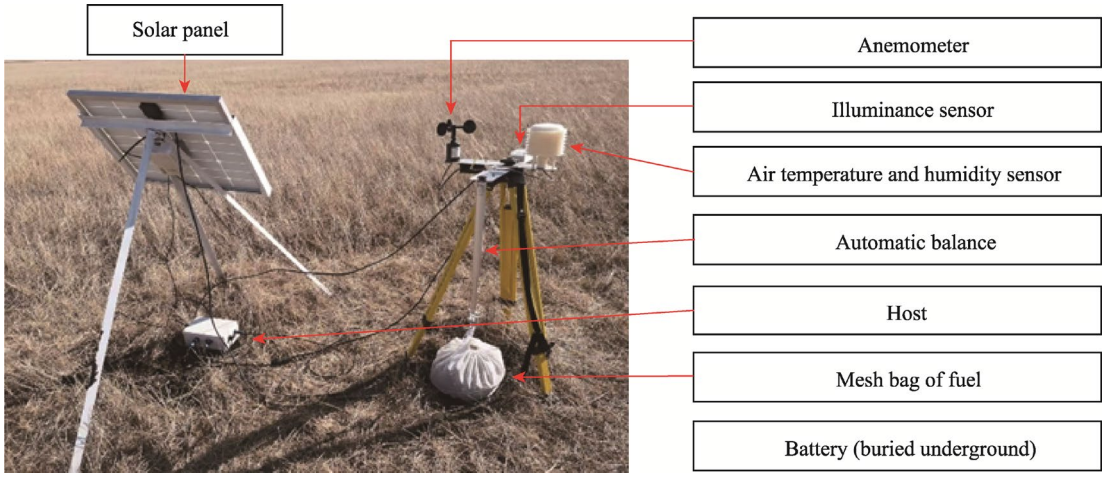

Fig. 2 Fuel moisture content meter and its components. The anemometer is placed 1 m above the ground.

2.3 Data collection

We monitored the real-time meteorological conditions and fuel weights at 2-h intervals from 9 April to 7 November in 2021. The dates before 9 April and after 7 November in 2021 were the midwinter period in the study area, which may damage the meters. In addition, the study area is always covered by a thick snow layer during the coldest period, which limits the availability of fuel on the surface. The 3 meters generated 7633 pieces of data in total, which we obtained in a TXT file format on the phone application. However, there were many anomalous values shown in the TXT file that were caused by abnormal meteorological conditions and issues with the weighted levers. After deleting anomalous values, only 3379 valid monitoring data points remained.

First, we calculated the DFMC (%) based on the wet weight and dry weight according to the

following equation (Slijepcevic et al., 2015):

$$DFMC = \frac{W_w - W_d}{W_d} \times 100\%, \quad (1)$$

where W_w is the wet weight (g) and W_d is the dry weight (g).

Then, to improve the DFMC prediction accuracy, it was necessary to build different models for different seasons. Because of the high-latitude location of Ergun City, we divided the monitoring period into four seasons according to temperature (Table 1) and in accordance with the published monograph (Chen, 2012) and built models for each season. The four seasons in Table 1 only include the sampling dates, not the entire year. Although the air in winter is very dry, the fuels still retain a certain level of moisture. According to local climate conditions, the fuels are covered by snow in the midwinter period, so there was no measurement during this period.

Table 1 Seasonal division of this study according to the average daily temperature

Season	Range of date	Temperature (T) threshold
Spring	15 May 2021–29 June 2021	10°C≤T<22°C
Summer	30 June 2021–31 July 2021	T≥22°C
Autumn	1 August 2021–9 October 2021	10°C≤T<22°C
Winter (apart from midwinter)	9 April 2021–14 May 2021 and 10 October 2021–7 November 2021	T<10°C

In addition, the meteorological variables provided by the DFMC meters were insufficient, so we downloaded the hourly variables of rainfall, atmospheric pressure, soil temperature, and soil humidity from the National Meteorological Science Data Centre of China (http://data.cma.cn/). Then, we obtained the DFMC values and eight meteorological and edaphic variables in real time. It is known that the DFMC is influenced not only by real-time meteorological conditions but also by past variables (Shmuel et al., 2022). Therefore, we calculated the time-lagged variables of each environmental variable. For every variable, we calculated the mean of the value at the current time and the value in the preceding 2 h as the 2 h time-lag variable, the mean of the current time value and the value in the preceding 4 h as the 4 h time-lag variable, and so on, until we calculated the 72 h time-lag variable. Each environmental variable had 36 derived lagged variables, and the total number of variables used for modelling was 296.

2.4 Data analysis

A variety of methods have been used to predict the DFMC (Lee et al., 2020; Dragozi et al., 2021; Fan and He, 2021). We chose three empirical models, including the boosted regression tree (BRT) model (Cai et al., 2012), extreme gradient boosting (XGB) model (Chen and Guestrin, 2016), and RF model (Fan and He, 2021), to predict the DFMC in a grassland of Ergun City. In this study, 70% of the data were used for training the models, and the remaining 30% were used for testing. We used RStudio 1.3 to perform the analysis and Origin 9.5 to draw the figures.

2.4.1 Boosted regression tree (BRT) model

The BRT model is a self-learning method based on the classified and regression tree algorithm. It generates multiple regression trees through random selection and a self-learning method, which can improve the stability and prediction accuracy of the model (De'ath, 2007; Elith et al., 2008). During the operation, a certain amount of data is randomly selected several times to analyse the influence of the independent variables on dependent variables, and the remaining data are used to test the fitting results. The BRT model has been widely used in ecological modelling (Cao et al., 2005; Li et al., 2014). In the BRT model, the n.tree parameter represents the total number of trees to fit; the shrinkage parameter is applied to each tree in the expansion (De'ath, 2007). The shrinkage parameter is also known as the learning rate or step-size reduction. After multiple tests,

the n.tree parameter was set to 1000, and the shrinkage parameter was set as 0.05.

2.4.2 Extreme gradient boosting (XGB) model

The XGB model, commonly known as the XGBoost model, is a scalable and end-to-end tree boosting system that proposes a novel sparsity-aware algorithm for sparse data and a weighted quantile sketch for approximate tree learning (Chen and Guestrin, 2016). It has the advantages of fast calculation speed, better goodness of fit, and superior processing of large-scale data. The XGB model has not only been applied in predicting the DFMC values (Shmuel et al., 2022) but has also been used for commercial sales prediction, customer behaviour prediction, product categorisation, motion detection, etc. (Chen and Guestrin, 2016). In the XGB model, the max_depth parameter refers to the maximum depth of the tree that needs to be set by users, with a common interval value of 3 to 10; the eta parameter refers to the shrinkage of the step size to prevent overfitting, and the common interval value is from 0.0 to 1.0. After multiple tests, the max_depth and eta parameters were set as 8 and 0.8, respectively.

2.4.3 Random forest (RF) model

The RF model is a kind of supervised machine learning algorithm based on decision trees. It gathers numerous classification trees to improve the prediction accuracy of the model. It is unnecessary to set the function form in advance, and this model can also overcome the complex interactions between covariates to obtain a high regression accuracy (Gao et al., 2020). The RF model has the advantages of higher accuracy than individual decision trees and lower sensitivity to parameter adjustment than other machine learning models (Su et al., 2020). In the RF model, the mtry parameter represents the number of variables used to split the tree at every node, the ntree parameter represents the number of decision trees, and the nodesize parameter represents the minimum number of nodes in the decision tree (Su et al., 2020). In this study, after parameter tuning, the mtry parameter defaulted to 3 and the nodesize parameter defaulted to 5 in the classification model. The ntree parameter was determined by multiple tests to determine how many decision trees would be obtained when the error in the model was relatively stable. After multiple tests, the ntree parameter was set as 500.

In addition, we performed variable importance analysis by deriving a variable importance measure, Increase in Node Purity (IncNodePurity), which can provide a method to assess the contribution of each predictor variable to the modelling performance. This is just a relative value and can be calculated using the decrease in tree node impurities attributable to each predictor variable (Su et al., 2020). A larger IncNodePurity indicates a stronger importance of these predictor variables (Karlson et al., 2015).

These models were built based on RStudio 1.3: the BRT model was performed using the gbm package; the XGB model was built using the xgboost package; and the RF model was built using the randomForest package. To evaluate the accuracy of these models, the mean absolute error (MAE) and R^2 of the model's prediction were calculated.

The MAE was calculated using the following expression:

$$\mathrm{MAE} = \frac{1}{n}\sum_{i=1}^{n}\left|\hat{y}_i - y_i\right|, \tag{2}$$

where n is the number of samples; i is an integer from 1 to n; y_i is the observed DFMC value; and $\hat{y}_i$ is the DFMC value predicted by the model (%).

R^2 can reflect the fitting degree of the regression line to the observed value, which can be calculated by the following equation:

$$R^2 = 1 - \frac{\sum_{i=1}^{n}\left(\hat{y}_i - y_i\right)^2}{\sum_{i=1}^{n}\left(\hat{y}_i - \overline{y}_i\right)^2}, \tag{3}$$

where $\overline{y}_i$ is the arithmetic mean of the observed DFMC values (%).

3 Results

3.1 The DFMC in various seasons

According to Table 1, we divided the valid monitoring values into 4 seasons, i.e., spring, summer, autumn, and winter, with 1273, 702, 645, and 759 observed values, respectively. The DFMC in the four seasons is shown in Figure 3. As we expected, the DFMC in summer was higher than that in other seasons, and winter had the lowest DFMC. This is consistent with the rainy summer and dry winter of the temperate continental monsoon climate.

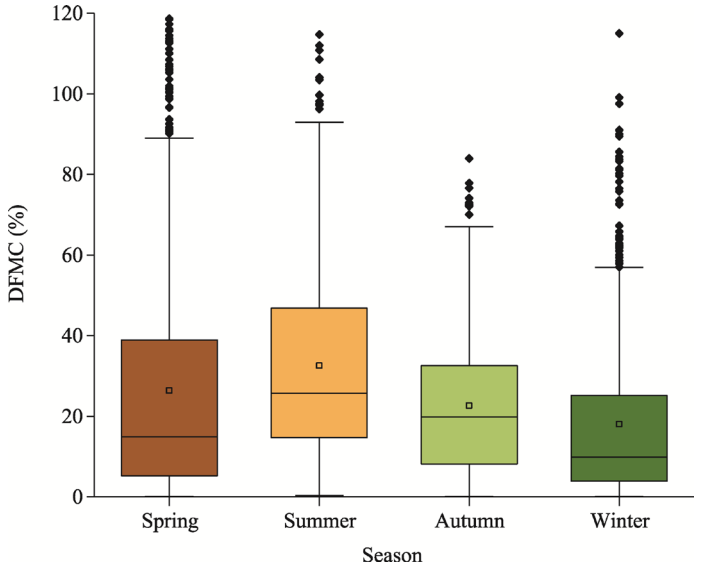

Fig. 3 Box plot of the seasonal differences in the dead fuel moisture content (DFMC). The upper and lower limits of the box indicate the 75th and 25th percentile values, respectively; the horizontal lines and small squares in each box represent the medians and means, respectively; the upper and lower whiskers show the maximum and minimum values, respectively; and the scattered points above the maximum values are outliers.

3.2 Model prediction accuracy

First, we built the three models, i.e., BRT, XGB, and RF models, based on all the training data, including the eight variables and their time-lag variables, and compared the accuracy according to the test data (Fig. 4). The RF model clearly not only achieved the highest accuracy, with an MAE score of 4.764%, but also had the best fitting effect, with an R^2 value of 0.847, whereas the XGB model showed inferior performance with an MAE score of 6.495% and a mediocre fitting effect with an R^2 value of 0.754, although its fitting line had the largest slope. The BRT model showed a larger MAE score of 7.709% and the worst fitting effect, with an R^2 value of 0.626.

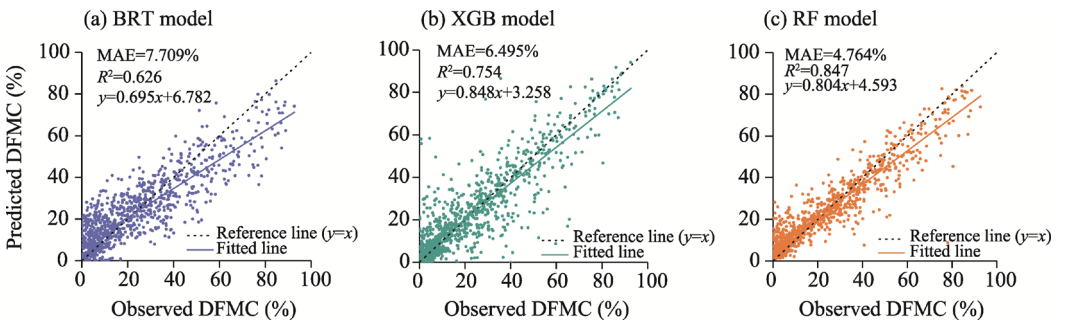

Fig. 4 Performances of the boosted regression tree (BRT) model (a), extreme gradient boosting (XGB) model (b), and random forest (RF) model (c) in predicting the DFMC. MAE is the mean absolute error.

We also built models for each season; the prediction accuracies and the performances are shown in Figure 5. We found that the RF model achieved the highest accuracy among the three models in every season, with MAE scores of 3.724%, 5.059%, 3.423%, and 3.873% for spring, summer, autumn, and winter, respectively. The R^2 values of the RF model were the highest in spring, summer, and autumn, with R^2 values of 0.922, 0.859, and 0.898, respectively. The XGB model seemed to perform worse than we expected because its R^2 values were the lowest in spring, summer, and autumn, although winter had an R^2 of 0.804, which was slightly higher than the 0.795 of the RF model. The BRT model showed moderate performance among the three models. In addition, the goodness of fit of the three models in spring and autumn was better than that in summer and winter.

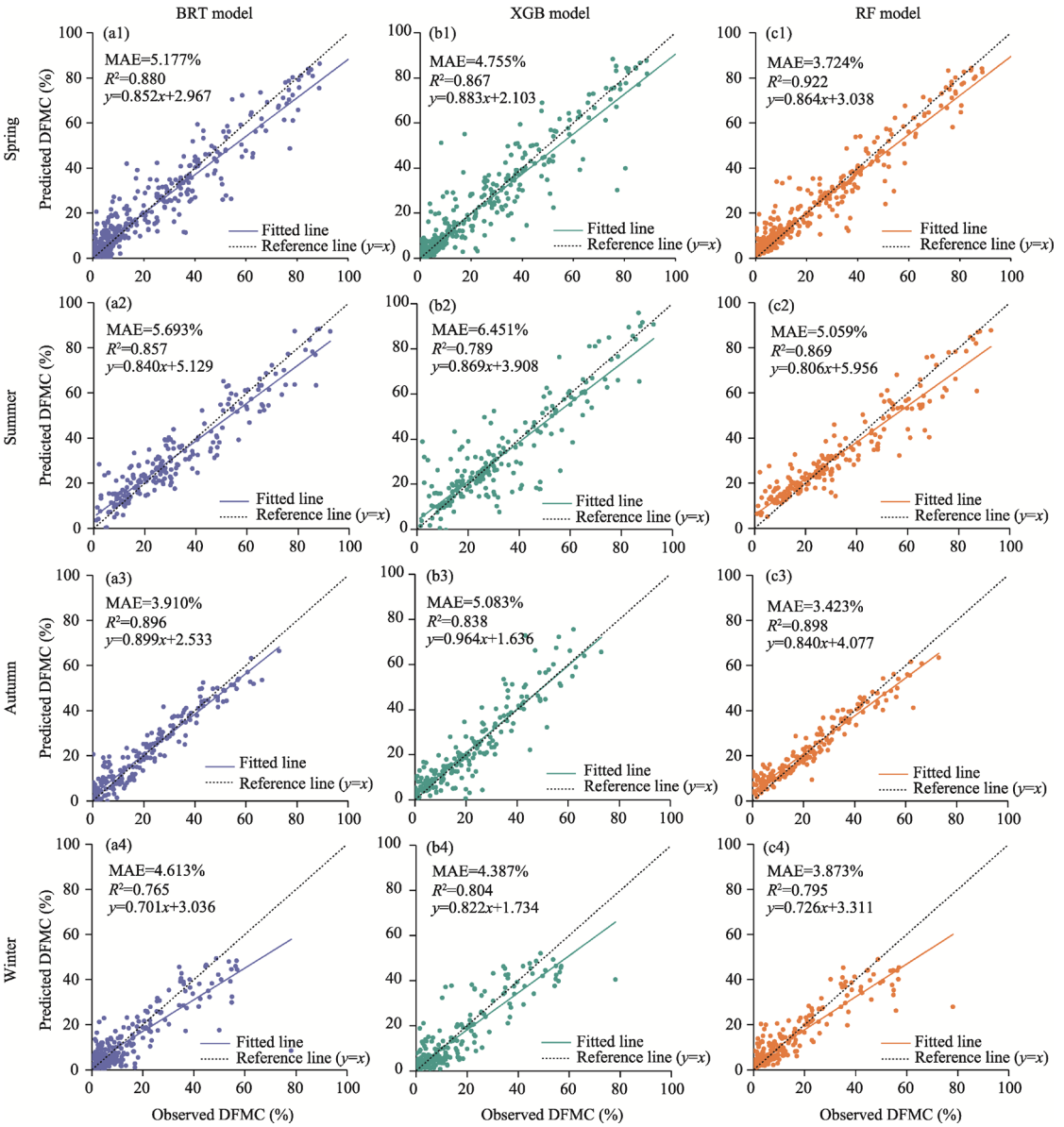

Fig. 5 Performance of the BRT, XGB, and RF models in spring (a1, b1, and c1), summer (a2, b2, and c2), autumn (a3, b3, and c3), and winter (a4, b4, and c4)

3.3 Variable importance

To find the key factors influencing the DFMC in each season, we drew variable importance line

charts with time lags according to the best fitting model, i.e., the RF model (Fig. 6). In spring, the soil humidity within –12 h (the negative numbers represent the hours before the time of measurement) had a certain influence on the DFMC, and a greater influence occurred within –4 h. In addition, rainfall showed a certain importance in two periods, from –50 to –40 h and from –22 to –10 h (Fig. 6a). In summer, the factors influencing the DFMC were complicated (Fig. 6b). The amount of rain had the highest importance in the period from –56 to –38 h. Soil humidity also showed a high importance, especially within –10 h. Relative air humidity had the highest influence at –6 h. In addition, the air pressure before –44 h showed a certain importance because rainfall was often accompanied by low air pressure. In autumn, the amount of rain had the highest importance in the period from –30 to –16 h. The change trend of the importance of soil humidity with a time lag was similar to that in both spring and summer. The importance of the relative air humidity increased periodically, reaching a peak at –14 h then decreasing rapidly to the lowest value at –6 h, and finally gradually increased until the time of measurement (Fig. 6c). In winter, the radiation at –26 h had the highest importance, which meant that the illumination intensity of the sun on the previous day was a key influencing factor on the DFMC. The importance of relative air humidity increased rapidly from –10 h and reached the highest value at –2 h. The influence of the amount of rain was relatively weak in winter but still showed a certain

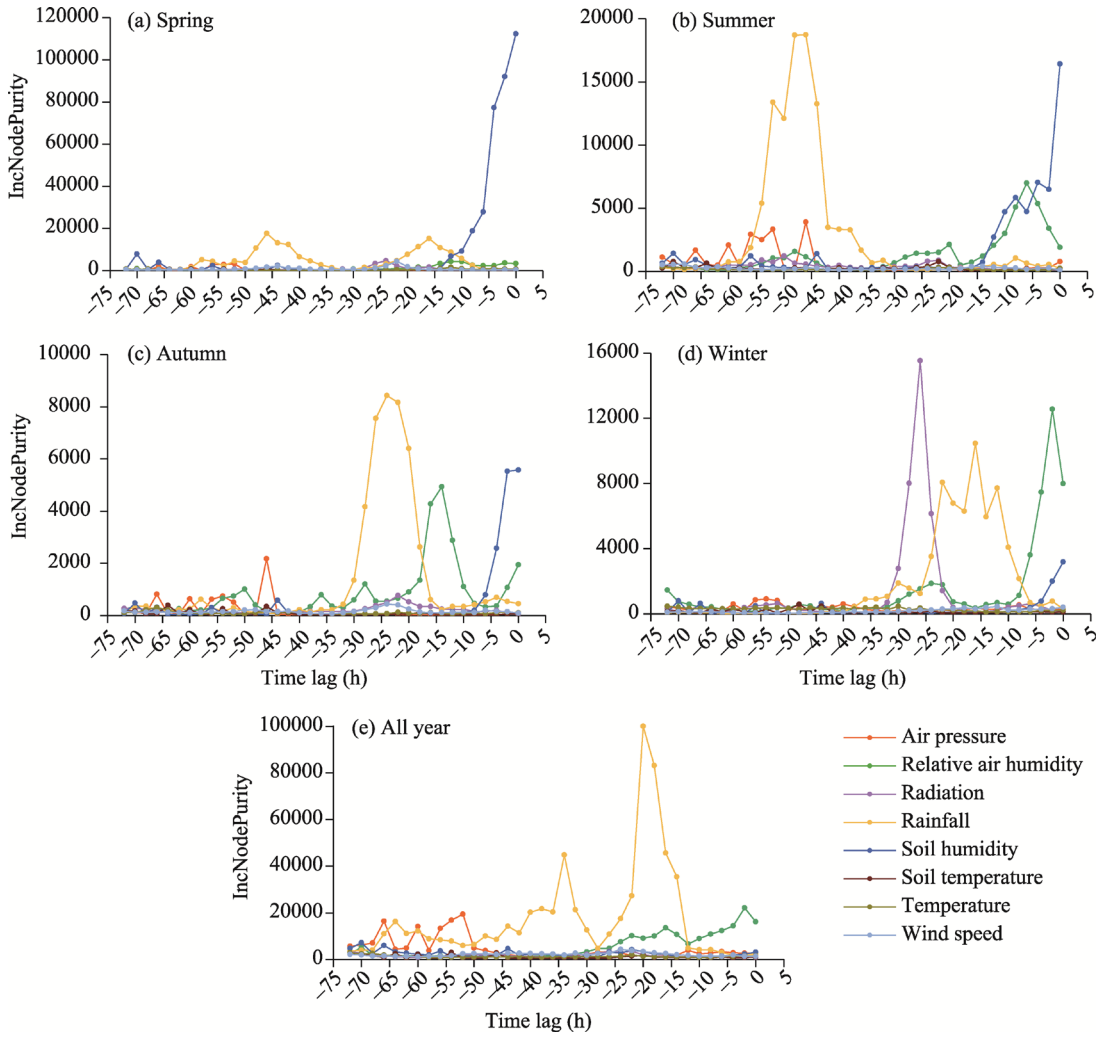

Fig. 6 Variable importance changes with time lags in the RF model in spring (a), summer (b), autumn (c), winter (d), and the all year (e). IncNodePurity is the degree of importance. The negative values of the *x*-axis represent the hours before the time of measurement.

importance in the period from –26 to –8 h (Fig. 6d). Overall, focusing on all data regardless of season, the amount of rain still showed the highest importance, with two peaks at –34 and –20 h. The relative air humidity and air pressure had some degree of importance (Fig. 6e).

3.4 Effect of time-lag length

We reran the RF-based model 37 times according to time-lag variables, with 2-h time-lagged variables removed each time, from the time of measurement to –72 h. Then, we calculated the MAE and R^2 based on the test data and plotted them against the time lags (Fig. 7). The results showed that except for in winter, the accuracies of the models incorporating time-lag variables at –2 h were obviously improved compared with those at the time of measurement, indicating that incorporating historical time-lag variables would significantly enhance the model prediction accuracy. In addition, the R^2 values increased and the MAE decreased with lengthening of the time lag. It is worth noting that the accuracies of the models significantly declined at approximately –44 h, while the most significant decline in the R^2 value and the most significant increase in the MAE occurred at approximately –24 h, implying that it is necessary to incorporate meteorological data within 1–2 d for the DFMC monitoring.

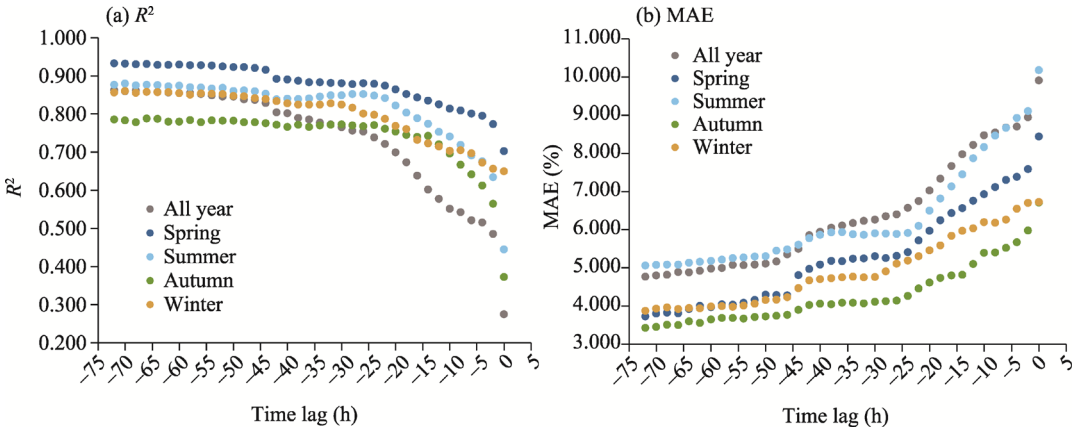

Fig. 7 Prediction performance along with the time-lag data 72 h before the time of measuring. (a), R^2; (b), MAE.

4 Discussion

4.1 DFMC prediction models

Building models for predicting the DFMC is crucial for forecasting the occurrences of grassland fires (Matthews, 2014). To accurately predict the DFMC, we built the BRT, XGB, and RF models for the whole monitoring period (Fig. 4) and for four seasons (Fig. 5) and compared their performances. The results indicated that the RF model performed the best, with a lower MAE and higher R^2, except in winter. Our results were consistent with those of previous studies (Capps et al., 2021; Masinda et al., 2021; Cunill Camprubí et al., 2022). Even though the XGB model showed excellent performance in the DFMC of forest modelling (Shmuel et al., 2022) and other fields (Chen and Guestrin, 2016), its performance was inadequate in this study. This may be due to the limited amount of monitoring data used, as building an XGB model requires a large amount of data (Chen and Guestrin, 2016). More data need to be collected in future research.

We also found that the seasonal models (Fig. 5) performed better than those for the whole monitoring period (Fig. 4), even though the summer models had larger MAE values. The larger errors in the summer models may be due to the intense rainfall in summer, which would have greatly influenced the DFMC. However, spring and autumn (not summer) are the high-risk seasons for fire (Hu et al., 2019) in this study area. Thus, predicting the DFMC in summer was less important. Our results proved that it is necessary to predict the DFMC for each season because the DFMC and environmental variables have various relationships in different seasons.

Therefore, we suggest building prediction models for the DFMC in spring and autumn to aid in grassland fire management planning in this study area. In addition, this study provides a method to predict the DFMC of grassland in North China, filling a research gap on the DFMC of grassland in China, but the DFMC and errors in this study were higher than those in other forestry areas (Fan and He, 2021; Shmuel et al., 2022). We propose that the dead fuels found on grasslands that lack tree canopies are more sensitive to rainfall than dead fuels in forests, as rainfall would markedly increase the DFMC.

4.2 Dominant variables influencing the DFMC

Recently, researchers have noted that there are time-lag issues that are relevant to the variables utilized for the DFMC prediction (Lopes et al., 2014; Yu et al., 2021). However, only a few studies have considered the time-lag variables of rainfall (Jin and Li, 2014) and relative air humidity (Zhang et al., 2015) in building the DFMC models. Other time-lagged variables, such as air pressure, soil humidity, radiation, air temperature, soil temperature, and wind speed, may also impact the DFMC. In this study, we comprehensively used all of these variables, as well as the associated variables derived by time lags, to build the DFMC models with improved accuracy. The best performing RF models were those built for the whole monitoring period and for the four seasons because different seasons with various weather patterns influence the DFMC (Hu et al., 2019). We used the IncNodePurity metric from the RF model to examine the variable importance in predicting the DFMC (Fig. 6). The importance of soil humidity in each season was clearly noted, which is consistent with previous studies (Qi et al., 2013; Rakhmatulina et al., 2021; Vinodkumar et al., 2021). The importance of soil humidity was more obvious in spring, which showed that the DFMC was more sensitive to soil humidity in the dry season. Rainfall in summer and autumn had dominant roles in controlling the DFMC and showed obvious hysteresis. This showed that rainfall had a continuous effect on the DFMC, with an approximately 1 d shorter effect in autumn and an approximately 2 d longer effect in summer. In winter, radiation showed the highest importance among the variables, and we believe that the sunny weather and better air quality in winter in the study area led to stronger radiation, thus reducing the DFMC in winter. However, in the all-year models, all of the variables except for the rain amount were less important. This also showed the importance of building models using time-lagged variables for every season to improve the model prediction accuracy (Hu et al., 2019).

4.3 Determination of the time span for time-lagged variables

In recent years, time-lag variables have been utilized to predict the DFMC (Jin and Li, 2014; Zhang et al., 2015). However, studies regarding the time spans are seldom documented. Some studies have focused on individual meteorological variables. According to the research conducted by Lee et al (2020), it is necessary to consider the rainfall before the time of measurement when predicting the DFMC. González et al. (2009) added a 2-h time-lag variable for air relative humidity to build empirical models for predicting dead fine fuel moisture. To explore the length of the time-lag variables that should be added, we calculated the goodness of fit and accuracy based on the RF model using time-lag variables of different lengths (Fig. 7). We found that the accuracy decreased with fluctuations from –44 h to the time of measurement, and the most obvious decrease occurred at –24 h. This was similar to the results of Shmuel et al. (2022), who stated that the model accuracy would be remarkably improved by adding time-lag variables of 20–30 h before the time of measurement. This indicates that the environmental conditions within 2 d are necessary and those within 1 d are significant for the DFMC prediction. Therefore, we suggest that 48 h time-lagged environmental variables, especially for rainfall, soil humidity, and relative air humidity, should be used in models to accurately estimate the DFMC of grassland.

4.4 Implications for grassland fuel management

Our results highlight the importance of 48 h time-lagged variables in predicting the DFMC of grassland and mapping grassland fire risks based on the DFMC, which will help identify high-priority areas for grassland fire monitoring. In addition, local grassland fire prevention

departments should pay more attention to monitoring environmental factors in spring and autumn. In Ergun City, there is a temperate continental monsoon season; in spring, the temperature quickly rises, more windy weather leads to dry air, and the soil moisture decreases. Hence, it is necessary to pay attention to the variation in soil moisture within 12 h (Fig. 6c). Similarly, the air in autumn is dry, and the soil moisture is low, so according to Figure 6c, it is necessary to monitor the variation in soil moisture within 8 h and the air relative humidity within 40 h. It is worth noting that there are differences between grassland fuels and forest fuels. For example, the thicknesses and sizes of the litter and the environmental conditions where the fuels are located are different, so their respective environmental variables will be different. Thus, there should also be differences in the management of forest and grassland fuels.

4.5 Limitations

Although we monitored the DFMC for nearly seven months and obtained a large amount of monitoring data, there were many anomalous values in the dataset that were attributed to the local weather conditions. This greatly reduced the amount of data, thus leading to discontinuities in the observations. Future research should be based on the continuously-monitored DFMC values and combine time series methods and the RF models to more precisely predict the DFMC. In addition, if the DFMC can be monitored for 3 to 4 years, its dynamic variation rule will be more clearly understood.

5 Conclusions

In this study, we used the DFMC data and the observed environmental variables from a grassland of Ergun City to build the DFMC prediction models for every season based on the BRT, XGB, and RF models. We found that the RF model performed better than the other two models, and the BRT and XGB models can be used as a reference to predict the DFMC in the Ergun study area. The RF model in spring and autumn had higher accuracy. Different seasons had various key influencing factors, and the degree of influence of these factors on the DFMC changed with time, revealing noticeable lags. For example, soil humidity at –10 h had a significant influence on the DFMC in spring; rainfall in summer exerted the highest influence at approximately –48 h, while it had the greatest influence in autumn at approximately –24 h; radiation in winter had a peak impact at –26 h. Adding the time-lag variables within 44 h clearly improved the fitting effect and prediction accuracy, indicating that environmental conditions within 48 h had a great influence on the DFMC. This study highlights the importance of considering 48 h time-lagged variables when predicting the DFMC of grassland and mapping grassland regions with fire risk based on the DFMC. This approach will help identify high-priority areas for grassland fire monitoring and prevention.

Acknowledgements

This study was funded by the National Key Research and Development Program of China Strategic International Cooperation in Science and Technology Innovation Program (2018YFE0207800) and the National Natural Science Foundation of China (31971483).

References

Bakšić N, Bakšić D, Jazbec A. 2017. Hourly fine fuel moisture model for *Pinus halepensis* (Mill.) litter. Agricultural and Forest Meteorology, 243: 93–99.

Bilgili E, Coskuner K A, Usta Y, et al. 2018. Modeling surface fuels moisture content in *Pinus brutia* stands. Journal of Forestry Research, 30(2): 577–587.

Cai W H, Yang J, Liu Z H, et al. 2012. Controls of post-fire tree recruitment in Great Xing'an Mountains in Heilongjiang Province. Acta Ecologica Sinica, 32(11): 3303–3312. (in Chinese)

Cao M C, Zhou G S, Weng E S. 2005. Application and comparison of generalized models and classification and regression tree in simulating tree species distribution. Acta Ecologica Sinica, 25(8): 2031–2040. (in Chinese)

Capps S B, Zhuang W, Liu R, et al. 2021. Modelling chamise fuel moisture content across California: A machine learning approach. International Journal of Wildland Fire, 31(2): 136–148.

Catchpole W R, Catchpole E A, Butler B W, et al. 1998. Rate of spread of free-burning fires in woody fuels in a wind Tunnel. Combustion Science and Technology, 131: 1–37.

Chen T Q, Guestrin C. 2016. XGBoost. In: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. San Francisco, USA.

Chen Y, Zhao L, Jiang Y D, et al. 2012. Division of Climate Season. Beijing: China Meteorological Press. (in Chinese)

Cunill Camprubí À, González-Moreno P, Resco de Dios V. 2022. Live fuel moisture content mapping in the Mediterranean Basin using random forests and combining MODIS spectral and thermal data. Remote Sensing, 14(13): 3162, doi: 10.3390/rs14133162.

Deak B, Valko O, Toeroek P, et al. 2014. Grassland fires in Hungary–experiences of nature conservationists on the effects of fire on biodiversity. Applied Ecology and Environmental Research, 12(1): 267–283.

De'ath G. 2007. Boosted trees for ecological modeling and prediction. Ecology, 88: 243–251.

Di Z L, Wu Y N, Song Y T, et al. 2019. Changes of extreme climate index in forest-steppe ecotone in Erguna. Chinese Journal of Ecology, 38(10): 3143–3152. (in Chinese)

Dragozi E, Giannaros T M, Kotroni V, et al. 2021. Dead fuel moisture content (DFMC) estimation using MODIS and meteorological data: the case of Greece. Remote Sensing, 13(21): 4224, doi: 10.3390/rs13214224.

Elith J, Leathwick J R, Hastie T. 2008. A working guide to boosted regression trees. Journal of Animal Ecology, 77(4): 802–813.

Fan C Q, He B B. 2021. A physics-guided deep learning model for 10-h dead fuel moisture content estimation. Forests, 12(7): 933, doi: 10.3390/rs13214224.

Fernandes P A M. 2001. Fire spread prediction in shrub fuels in Portugal. Forest Ecology and Management, 144(1–3): 67–74.

Fontenele H G V, Cruz-Lima L F S, Pacheco-Filho J L, et al. 2020. Burning grasses, poor seeds: post-fire reproduction of early-flowering Neotropical savanna grasses produces low-quality seeds. Plant Ecology, 221(12): 1265–1274.

Gao C, Lin H L, Hu H Q, et al. 2020. A review of models of forest fire occurrence prediction in China. The Journal of Applied Ecology, 31(9): 3227–3240. (in Chinese)

González A D R, Hidalgo J A V, González J G Á. 2009. Construction of empirical models for predicting *Pinus sp*. dead fine fuel moisture in NW Spain. I: Response to changes in temperature and relative humidity. International Journal of Wildland Fire, 18(1): 71–83.

Hiers J K, Stauhammer C L, O'Brien J J, et al. 2019. Fine dead fuel moisture shows complex lagged responses to environmental conditions in a saw palmetto (*Serenoa repens*) flatwoods. Agricultural and Forest Meteorology, 266–267: 20–28.

Hu H Q, Luo B Z, Luo S S, et al. 2019. Water content of surface ground fuel in *Larix gmelinii-Betula platyphylla* mixed forest of Nanwenhe, Daxing'an Mountains. Chinese Journal of Ecology, 38(5): 1314–1321. (in Chinese)

Jin S, Li J M. 2014. Prediction on moisture contents of typical forest dead combustible fuels of an ecotones in Qingan county of Heilongjiang province. Journal of Central South University of Forestry & Technology, 34(12): 27–34. (in Chinese)

Karlson M, Ostwald M, Reese H, et al. 2015. Mapping tree canopy cover and aboveground biomass in Sudano-Sahelian woodlands using Landsat 8 and random forest. Remote Sensing, 7(8): 10017–10041.

Lee H, Won M, Yoon S, et al. 2020. Estimation of 10-hour fuel moisture content using meteorological data: a model inter-comparison study. Forests, 11: 982, doi: 10.3390/f11090982.

Lei W D, Yu Y, Li X H, et al. 2022. Estimating dead fine fuel moisture content of forest surface, based on wireless sensor network and back-propagation neural network. International Journal of Wildland Fire, 31(4): 369–378.

Li C L, Liu M, Hu Y M, et al. 2014. Driving forces analysis of urban expansion based on boosted regression trees and Logistic regression. Acta Ecologica Sinica, 34(3): 727–737. (in Chinese)

Limb R F, Fuhlendorf S D, Engle D M, et al. 2011. Pyric–herbivory and cattle performance in grassland ecosystems. Rangeland Ecology & Management, 64(6): 659–663.

Lopes S, Viegas D X, Teixeira de Lemos L, et al. 2014. Equilibrium moisture content and timelag of dead *Pinus pinaster* needles. International Journal of Wildland Fire, 23(5): 721–732.

Man Z Y, Hu H Q, Zhang Y L, et al. 2019. Dynamic change and prediction model of moisture content of surface fuel in Maoer Mountain of northeastern China. Journal of Beijing Forestry University, 41(3): 49–57. (in Chinese)

Masinda M M, Li F, Liu Q, et al. 2021. Prediction model of moisture content of dead fine fuel in forest plantations on Maoer Mountain, Northeast China. Journal of Forestry Research, 32(5): 2023–2035.

Masinda M M, Li F, Qi L, et al. 2022. Forest fire risk estimation in a typical temperate forest in Northeastern China using the Canadian forest fire weather index: Case study in autumn 2019 and 2020. Natural Hazards, 111: 1085–1101.

Matthews S, Gould J, McCaw L. 2010. Simple models for predicting dead fuel moisture in eucalyptus forests. International Journal of Wildland Fire, 19(4): 459–467.

Matthews S. 2014. Dead fuel moisture research: 1991–2012. International Journal of Wildland Fire, 23(1): 78–92.

Mouillot F, Field C B. 2005. Fire history and the global carbon budget: a 1°×1° fire history reconstruction for the 20th century. Global Change Biology, 11(3): 398–420.

Muro J, Linstädter A, Magdon P, et al. 2022. Predicting plant biomass and species richness in temperate grasslands across regions, time, and land management with remote sensing and deep learning. Remote Sensing of Environment, 282: 113262, doi: 10.1016/j.rse.2022.113262.

Nelson R M. 2000. Prediction of diurnal change in 10-h fuel stick moisture content. Canadian Journal of Forest Research-Revue Canadienne De Recherche Forestiere, 30(7): 1071–1087.

Nieto H, Aguado I, Chuvieco E, et al. 2010. Dead fuel moisture estimation with MSG–SEVIRI data. Retrieval of meteorological data for the calculation of the equilibrium moisture content. Agricultural and Forest Meteorology, 150(7–8): 861–870.

Petermann J S, Buzhdygan O Y. 2021. Grassland biodiversity. Current Biology, 31(19): 1195–1201.

Podur J, Martell D L, Csillag F. 2003. Spatial patterns of lightning-caused forest fires in Ontario, 1976–1998. Ecological Modelling, 164(1): 1–20.

Qi Y, Dennison P E, Spencer J, et al. 2013. Monitoring live fuel moisture using soil moisture and remote sensing proxies. Fire Ecology, 8(3): 71–87.

Rakhmatulina E, Stephens S, Thompson S. 2021. Soil moisture influences on Sierra Nevada dead fuel moisture content and fire risks. Forest Ecology and Management, 496: 119379, doi: 10.1016/j.foreco.2021.119379.

Resco de Dios V, Fellows A W, Nolan R H, et al. 2015. A semi-mechanistic model for predicting the moisture content of fine litter. Agricultural and Forest Meteorology, 203: 64–73.

Resco de Dios V, Hedo J, Cunill Camprubí À, et al. 2021. Climate change induced declines in fuel moisture may turn currently fire-free Pyrenean mountain forests into fire-prone ecosystems. Science of the Total Environment, 797: 149104, doi: 10.1016/j.scitotenv.2021.149104.

Schunk C, Wastl C, Leuchner M, et al. 2017. Fine fuel moisture for site- and species-specific fire danger assessment in comparison to fire danger indices. Agricultural and Forest Meteorology, 234–235: 31–47.

Sharma S, Carlson J D, Krueger E S, et al. 2021. Soil moisture as an indicator of growing-season herbaceous fuel moisture and curing rate in grasslands. International Journal of Wildland Fire, 30(1): 57–69.

Shmuel A, Ziv Y, Heifetz E. 2022. Machine-Learning-based evaluation of the time-lagged effect of meteorological factors on 10-hour dead fuel moisture content. Forest Ecology and Management, 505: 119897, doi: 10.1016/j.foreco.2021.119897.

Slijepcevic A, Anderson W R, Matthews S, et al. 2015. Evaluating models to predict daily fine fuel moisture content in eucalypt forest. Forest Ecology and Management, 335: 261–269.

Stocks B J, Lawson B D, Alexander M E, et al. 1989. Canadian forest fire danger rating system - an overview. Forestry Chronicle, 65(4): 258–265.

Su H Y, Shen W J, Wang J R, et al. 2020. Machine learning and geostatistical approaches for estimating aboveground biomass in Chinese subtropical forests. Forest Ecosystems, 7(1): 64, doi: 10.1186/s40663-020-00276-7.

Sun L, Liu Q, Hu T X. 2021. Advances in research on prediction model of moisture content of surface dead fuel in forests. Scientia Silvae Sinicae, 57(4): 142–152. (in Chinese)

Vinodkumar V, Dharssi I, Yebra M, et al. 2021. Continental-scale prediction of live fuel moisture content using soil moisture information. Agricultural and Forest Meteorology, 307: 108503, doi: 10.1016/j.agrformet.2021.108503.

Wilson R A. 1985. Observations of extinction and marginal burning states in free burning porous fuel beds. Combustion Science and Technology, 44(3–4): 179–193.

Xing J J, Qu Z L. 2017. Ground surface fuel moisture content by mixed effects models in Daxing'an Mountains. Journal of North-East Forestry University, 45(3): 58–62. (in Chinese)

Yebra M, Chuvieco E, Riano D. 2008. Estimation of live fuel moisture content from MODIS images for fire risk assessment. Agricultural and Forest Meteorology, 148(4): 523–536.

Yu H Z, Shu L F, Yang G, et al. 2021. Comparison of vapour-exchange methods for predicting hourly twig fuel moisture contents of larch and birch stands in the Daxinganling Region, China. International Journal of Wildland Fire, 30(6): 462–466.

Zhang Y L, Zhang H, Jin S. 2015. Effects of season change and rainfall on forecast model accuracy of predicting fine fuels in forests in Pangu Forest Farm. Journal of Central South University of Forestry & Technology, 35(8): 5–12. (in Chinese)

Zhu L J, Webb G I, Yebra M, et al. 2021. Live fuel moisture content estimation from MODIS: A deep learning approach. ISPRS Journal of Photogrammetry and Remote Sensing, 179: 81–91.